

Seq ID NO: 7

Database : A_Geneseq_23Sep04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	%		Query			Description
	Score	Match	Length	DB	ID	
1	2794	100.0	497	3	AA93750	Aay93750 Amino aci <i>run work</i>
2	439.5	15.7	174	6	AAE30346	Aae30346 Perna can
3	439.5	15.7	175	6	AAE30347	Aae30347 Crassostr
4	260	9.3	1529	2	AAR97985	Aar97985 CORK pota
5	217	7.8	351	2	AAR24393	Aar24393 Sequence
6	178	6.4	339	6	ADA35264	Ada35264 Acinetoba
7	173.5	6.2	244	2	AAR67409	Aar67409 Rat super
8	173.5	6.2	244	5	AAM52476	Aam52476 Superoxid
9	173.5	6.2	244	7	ADD48518	Add48518 Rat Prote
10	172.5	6.2	221	2	AAR27934	Aar27934 GAG fusio

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	% Query		Match	Length	DB	ID	Description
	Score	Match					
1	178	6.4	339	4	US-09-328-352-6551	Sequence 6551, Ap	
2	173.5	6.2	244	3	US-08-679-493A-188	Sequence 188, App	
3	168	6.0	150	2	US-08-722-050-9	Sequence 9, Appli	
4	168	6.0	150	4	US-09-883-985-9	Sequence 9, Appli	
5	167	6.0	154	3	US-08-679-493A-211	Sequence 211, App	
6	166	5.9	151	2	US-08-722-050-10	Sequence 10, Appl	
7	166	5.9	151	4	US-09-883-985-10	Sequence 10, Appl	
8	165.5	5.9	152	2	US-08-722-050-12	Sequence 12, Appl	
9	165.5	5.9	152	4	US-09-883-985-12	Sequence 12, Appl	
10	164.5	5.9	153	3	US-08-679-493A-207	Sequence 207, App	
11	164	5.9	151	3	US-08-679-493A-191	Sequence 191, App	
12	163.5	5.9	153	3	US-08-679-493A-201	Sequence 201, App	
13	161.5	5.8	153	3	US-08-679-493A-202	Sequence 202, App	
14	160.5	5.7	152	6	5171680-3	Patent No. 5171680	
15	160	5.7	1099	4	US-09-881-654-4	Sequence 4, Appli	
16	160	5.7	1099	4	US-10-637-323-4	Sequence 4, Appli	
17	159.5	5.7	699	4	US-09-538-092-995	Sequence 995, App	
18	159	5.7	166	3	US-08-679-493A-209	Sequence 209, App	

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%					
Result		Query					
No.	Score	Match	Length	DB	ID	Description	
1	170	6.1	152	17	US-10-425-115-233754	Sequence 233754,	
2	170	6.1	153	15	US-10-425-114-48136	Sequence 48136, A	
3	170	6.1	153	15	US-10-425-114-52073	Sequence 52073, A	
4	170	6.1	153	15	US-10-425-114-52143	Sequence 52143, A	
5	170	6.1	153	15	US-10-425-114-59106	Sequence 59106, A	
6	170	6.1	153	15	US-10-425-114-61368	Sequence 61368, A	
7	170	6.1	153	15	US-10-425-114-62898	Sequence 62898, A	
8	170	6.1	153	15	US-10-425-114-66160	Sequence 66160, A	
9	170	6.1	153	15	US-10-425-114-72460	Sequence 72460, A	

Database : PIR_79:*
 1: pir1:*
 2: pir2:*
 3: pir3:*
 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	% Query		Match Length	DB	ID	Description
	Score	Match				
1	213	7.6	351	1	KGZQHL	histidine-rich gly
2	204.5	7.3	735	2	T45059	hypothetical prote
3	178	6.4	152	2	JW0084	superoxide dismuta
4	178	6.4	852	2	A34373	histidine-rich cal
5	174.5	6.2	251	2	S52859	superoxide dismuta
6	173.5	6.2	152	2	T06570	superoxide dismuta
7	173.5	6.2	244	2	A49097	superoxide dismuta
8	173	6.2	1840	2	T29091	transitin - chicke
9	168	6.0	151	2	A29077	superoxide dismuta
10	167	6.0	154	1	DSBYC	superoxide dismuta
11	164.5	5.9	154	1	DSHOCZ	superoxide dismuta
12	164	5.9	152	2	S07007	superoxide dismuta
13	163	5.8	152	2	S22508	superoxide dismuta
14	163	5.8	152	2	S72235	superoxide dismuta

Database : UniProt_02:*
 1: uniprot_sprot:*
 2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	2790	99.9	517	2	Q9BKB9	Q9bkb9 perna canal
2	439.5	15.7	174	2	Q86FW9	Q86fw9 crassostrea
3	221	7.9	294	2	Q7QDP9	Q7qdp9 anopheles g
4	213	7.6	351	1	HRPX_PLALO	P04929 plasmodium
5	204.5	7.3	735	2	Q9NES7	Q9nes7 caenorhabdi
6	196.5	7.0	2245	2	Q8IAM6	Q8iam6 plasmodium
7	191	6.8	722	2	Q7YS21	Q7ys21 macaca fasc
8	178.5	6.4	726	2	Q9QZV4	Q9qzv4 mus musculu
9	178	6.4	152	1	SODC_SOYBN	Q7mlr5 glycine max
10	178	6.4	852	1	SRCH_RABIT	P16230 oryctolagus
11	177	6.3	738	2	Q9WVE4	Q9wve4 mus musculu
12	175	6.3	151	1	SODC_HALRO	P81926 halocynthia
13	174.5	6.2	251	2	Q64466	Q64466 mus musculu
14	174	6.2	152	2	Q9ZNQ4	Q9znq4 cicer ariet
15	173.5	6.2	151	1	SODC_PEA	Q02610 pisum sativ

RESULT 1

Q9BKB9

ID Q9BKB9 PRELIMINARY; PRT; 517 AA.
 AC Q9BKB9;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Pernin precursor.
 OS Perna canaliculus (greenshell mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Mytiloida;
 OC Mytiloidea; Mytilidae; Perna.
 OX NCBI_TaxID=38949;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21186417; PubMed=11290459;
 RA Scotti P.D., Dearing S.C., Greenwood D.R., Newcomb R.D.;
 RT "Pernin: a novel self-aggregating haemolymph protein from the New
 RT Zealand green-lipped mussel Perna canaliculus (bivalvia: mytilidae).";
 RL Comp. Biochem. Physiol. B, Biochem. Mol. Biol. 128:767-779(2001).
 DR EMBL; AF273766; AAK20952.1; -.
 DR HSSP; P00445; 1F1G.
 DR GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0006801; P:superoxide metabolism; IEA.
 DR InterPro; IPR001424; SOD_CU_ZN.
 DR Pfam; PF00080; Sod_Cu; 3.
 DR PRINTS; PR00068; CUZNDISMTASE.
 KW Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 517 pernin.
 SQ SEQUENCE 517 AA; 57222 MW; 87B8FBFFE855501E CRC64;

Query Match 99.9%; Score 2790; DB 2; Length 517;
 Best Local Similarity 99.8%; Pred. No. 1.6e-198;
 Matches 496; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	DGEQCNDGQNKDDHDDHDDHDDHDDDDDETMHYAQCEMEPNPHMASSLHHHVHGSIEL	60
Db	21	DGEQCNDGQNKDDHDDHDDHDDHDDDDDETMHYAQCEMEPNPHMASSLHHHVHGSIEL	80
Qy	61	SQKGHGAVYLELHLVGFNTSEDHDDHHHGLHLHMLGDMSAGCDSIGELYNAHPEKHADPG	120
Db	81	SQQGHGAVYLELHLVGFNTSEDHDDHHHGLHLHMLGDMSAGCDSIGELYNAHPEKHADPG	140
Qy	121	DLGDLVDDDRGVVNEVHHYAWLDIDGTAPNTEALIGHSMITLQGSHTDADTPASRIACCV	180
Db	141	DLGDLVDDDRGVVNEVHHYAWLDIDGTAPNTEALIGHSMITLQGSHTDADTPASRIACCV	200
Qy	181	IGHGKARPETAAALHHELEEDKTEHYAHCDVRSNTHQPKALHHHVHGTIDFKQVGYGDLE	240
Db	201	IGHGKARPETAAALHHELEEDKTEHYAHCDVRSNTHQPKALHHHVHGTIDFKQVGYGDLE	260
Qy	241	VSYHLEGFNVSDDHKDLHLDVQIYANGDLTSGCDNLGAKYDPHEDYHSELGDLGDIHDDD	300
Db	261	VSYHLEGFNVSDDHKDLHLDVQIYANGDLTSGCDNLGAKYDPHEDYHSELGDLGDIHDDD	320
Qy	301	HGVVNESHRYSWINIFGDDSVLGRSIAIHQRDLHKSAKIACCVIGRGQSHPEIVHRAKC	360

Db	321 HGVVNESHRYSWINIFGDDSVLGRSIAIHQRDHLHKSAKIACCVIGRGQSHPEIVHRACK	380	
Qy	361 VVRPNTESTGLHHHVSGSITFEQTPGGSTHMTADLKGFNVSEDLSHHRHGVQLHEWGDMS	420	
Db	381 VVRPNTESTGLHHHVSGSITFEQTPGGSTHMTADLKGFNVSEDLSHHRHGVQLHEWGDMS	440	
Qy	421 HGCHSLGRMYGHGDDAHDPKRPGDLGDVIDDSHGIVHSTRTFDHLNVEDLNARSLVIMQG	480	
Db	441 HGCHSLGRMYGHGDDAHDPKRPGDLGDVIDDSHGIVHSTRTFDHLNVEDLNARSLVIMQG	500	
Qy	481 GHEVESERVACCVIGRA	497	
Db	501 GHEVESERVACCVIGRA	517	

See ID 1026

Database : EST:*
 1: gb_est1:*
 2: gb_est2:*
 3: gb_htc:*
 4: gb_est3:*
 5: gb_est4:*
 6: gb_est5:*
 7: gb_est6:*
 8: gb_gss1:*
 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%							
Result	Query						
No.	Score	Match	Length	DB	ID	Description	
1	155.8	10.4	688	6	CD649186	CD649186	AUF_104_N
2	149.4	10.0	704	6	CD648295	CD648295	AUF_102_G
3	147.8	9.9	682	6	CD648076	CD648076	AUF_101_M
4	147.8	9.9	697	6	CD647088	CD647088	AUF_107_A
5	147.8	9.9	697	6	CD647705	CD647705	AUF_108_L
6	147.8	9.9	706	6	CD649879	CD649879	CvGil0058
7	147	9.9	698	6	CD650428	CD650428	CvGil0113
8	146.2	9.8	696	6	CD648647	CD648647	AUF_103_F
9	146.2	9.8	699	6	CD648443	CD648443	AUF_102_M
10	146.2	9.8	720	6	CD648998	CD648998	AUF_104_E
11	146.2	9.8	725	6	CD649188	CD649188	AUF_104_N
12	146.2	9.8	734	6	CD648621	CD648621	AUF_103_E
13	145.4	9.8	696	6	CD648155	CD648155	AUF_101_P
14	145.4	9.8	713	6	CD649071	CD649071	AUF_104_I
15	144.6	9.7	698	6	CD648763	CD648763	AUF_103_K

Scoring table: OLIGO_NUC
 Gapop 60.0 , Gapext 60.0

Searched: 32822875 seqs, 18219865908 residues

Word size : 0

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_htc:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gss1:*
- 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	26	1.7 766	8	BH948315	BH948315 obu82g07.
2	26	1.7 1101	9	CNS00HD3	AL073332 Drosophil
3	25	1.7 529	5	BQ118156	BQ118156 EST603732
4	25	1.7 756	6	CB942058	CB942058 AGENCOURT
5	25	1.7 946	6	CF265550	CF265550 AGENCOURT
6	23	1.5 541	1	AI724181	AI724181 RHIZ1_8_B
7	23	1.5 602	5	BW326909	BW326909 BW326909

Database : GenEmbl:*
 1: gb_ba:*
 2: gb_htg:*
 3: gb_in:*
 4: gb_om:*
 5: gb_ov:*
 6: gb_pat:*
 7: gb_ph:*
 8: gb_pl:*
 9: gb_pr:*
 10: gb_ro:*
 11: gb_sts:*
 12: gb_sy:*
 13: gb_un:*
 14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result			%		Query		DB	ID	Description
	No.	Score	Match	Length	Match	Length			
	1	1490.6	100.0	1491	6	BD268169			BD268169 Serine pr
	2	1490.6	100.0	1611	6	BD268170			BD268170 Serine pr
	3	1484.2	99.5	1700	3	AF273766			AF273766 Perna can
	4	128.4	8.6	603	3	AY256853			AY256853 Crassostr
c	5	88.4	5.9	115758	9	AC104634			AC104634 Homo sapi
	6	86	5.8	110000	2	PFMAL13_24			Continuation (25 o
c	7	74.8	5.0	164347	9	AC104805			AC104805 Homo sapi
	8	74.8	5.0	186278	9	AC079176			AC079176 Homo sapi
c	9	74	5.0	75111	5	BX276082			BX276082 Zebrafish

RESULT 1

BD268169

LOCUS BD268169 1491 bp DNA linear PAT 17-JUL-2003

DEFINITION Serine protease inhibitors.

ACCESSION BD268169

VERSION BD268169.1 GI:33077937

KEYWORDS JP 2002534063-A/1.

SOURCE unidentified

ORGANISM unidentified

unclassified.

REFERENCE 1 (bases 1 to 1491)

AUTHORS Scotti,P.D., Dearing,S.C., Greenwood,D.R. and Newcomb,R.D.

TITLE Serine protease inhibitors

JOURNAL Patent: JP 2002534063-A 1 15-OCT-2002;

THE HORTICULTURE AND FOOD RESEARCH INSTITUTE OF NEW ZEALAND LTD

COMMENT OS Shellfish

PN JP 2002534063-A/1

PD 15-OCT-2002

PF 23-DEC-1999 JP 2000591076

PR 23-DEC-1998 NZ 333568,23-JUL-1999 NZ 336906 PI

PAUL DOUGLAS SCOTTI,SALLY CAROLINE DEARING,DAVID ROGER PI

GREENWOOD,

PI RICHARD DAVID NEWCOMB

PC C12N15/09,A23L1/305,A61K38/00,A61P7/04,A61P43/00,C07K1/14, PC
C07K14/435,

PC C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N9/99// (C12N9/99,C12R1:
91),

PC C12N15/00,C12N5/00,A61K37/02

CC Serine protease inhibitors

FH Key Location/Qualifiers

FT source 1. .1491

FT /organism='Shellfish'.

FEATURES Location/Qualifiers

source 1. .1491

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 1490.6; DB 6; Length 1491;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1491; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 GAYGGGGAGCAGTGTAACGATGGGCAGAACAAAGATGACCACCATGACGACCACCACGAT 60
          |||
Db      1 GAYGGGGAGCAGTGTAACGATGGGCAGAACAAAGATGACCACCATGACGACCACCACGAT 60

Qy     61 GATCACCATGACGACCATGATGATGATGATGAAACAATGCACTATGCCCAGTGTGAAATG 120
          |||
Db     61 GATCACCATGACGACCATGATGATGATGATGAAACAATGCACTATGCCCAGTGTGAAATG 120

Qy    121 GAACCAAACCCTCATATGGCTAGCAGCCTTCACCACCATGTCCATGGCAGCATAGAGTTG 180
          |||
Db    121 GAACCAAACCCTCATATGGCTAGCAGCCTTCACCACCATGTCCATGGCAGCATAGAGTTG 180

Qy    181 TCACAGAAGGGTCATGGAGCTGTTTATCTAGAACTTCATCTTGTCGGATTCAACACAAGT 240
          |||

```

Db 181 TCACAGAAGGGTCATGGAGCTGTTTATCTAGAACTTCATCTTGTCTGGATTCAACACAAGT 240
 Qy 241 GAAGACCATGACGACCACCATCATGGACTTCATCTGCACATGCTTGGTGACATGTCAGCA 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 GAAGACCATGACGACCACCATCATGGACTTCATCTGCACATGCTTGGTGACATGTCAGCA 300
 Qy 301 GGTGTGATTCTATTGGCGAACTGTACAATGCTCACCCAGAAAAACATGCTGACCCTGGT 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 GGTGTGATTCTATTGGCGAACTGTACAATGCTCACCCAGAAAAACATGCTGACCCTGGT 360
 Qy 361 GACCTCGGTGACCTGGTTGACGATGATAGGGGCGTGGTTAATGAAGTTCATCATTATGCT 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 GACCTCGGTGACCTGGTTGACGATGATAGGGGCGTGGTTAATGAAGTTCATCATTATGCT 420
 Qy 421 TGGTTGGACATTGATGGTACAGCACCAAACACCGAAGCTCTCATTGGACACTCAATGACT 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 421 TGGTTGGACATTGATGGTACAGCACCAAACACCGAAGCTCTCATTGGACACTCAATGACT 480
 Qy 481 ATTTTACAAGGGAGTCACACCGATGCTGATACCCAGCCAGTAGAATCGCCTGTTGTGTT 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 481 ATTTTACAAGGGAGTCACACCGATGCTGATACCCAGCCAGTAGAATCGCCTGTTGTGTT 540
 Qy 541 ATTGGTCATGGAAAAGCTCGCCCAGAAACAGCAGCTGCTCTACATCACGAGCTAGAGGAA 600
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 541 ATTGGTCATGGAAAAGCTCGCCCAGAAACAGCAGCTGCTCTACATCACGAGCTAGAGGAA 600
 Qy 601 GATAAACTGAGCATTATGCCCATTGTGACGTAAGATCTAATACACACCAACCAAAGGCT 660
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 601 GATAAACTGAGCATTATGCCCATTGTGACGTAAGATCTAATACACACCAACCAAAGGCT 660
 Qy 661 CTTCATCATCATGTCCACGGAACCATCGATTTCAAACAAGTTGGTTATGGTGACCTTGAA 720
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 661 CTTCATCATCATGTCCACGGAACCATCGATTTCAAACAAGTTGGTTATGGTGACCTTGAA 720
 Qy 721 GTGTCCTACCATTTAGAGGGATTTAATGTAAGTGATGACCACAAAGATCATCTCCATGAC 780
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 721 GTGTCCTACCATTTAGAGGGATTTAATGTAAGTGATGACCACAAAGATCATCTCCATGAC 780
 Qy 781 GTACAGATCTACGCCAACGGTGACCTGACCAGTGGATGTGATAACCTCGGTGCTAAATAT 840
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 781 GTACAGATCTACGCCAACGGTGACCTGACCAGTGGATGTGATAACCTCGGTGCTAAATAT 840
 Qy 841 GATCCTCATGAAGATTACCACAGTGAGTTGGGTGATCTAGGAGATATTCACGATGATGAC 900
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 841 GATCCTCATGAAGATTACCACAGTGAGTTGGGTGATCTAGGAGATATTCACGATGATGAC 900
 Qy 901 CATGGCGTTGTCAATGAAAGCCACAGATATTCTTGGATCAATATCTTCGGTGATGACAGT 960
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 901 CATGGCGTTGTCAATGAAAGCCACAGATATTCTTGGATCAATATCTTCGGTGATGACAGT 960
 Qy 961 GTCCTGGGACGTTCTATTGCCATTACCAAAGAGACCATCTTCATAAAAGTGCCAAAATT 1020
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 961 GTCCTGGGACGTTCTATTGCCATTACCAAAGAGACCATCTTCATAAAAGTGCCAAAATT 1020
 Qy 1021 GCCTGTTGTGTCATAGGACGTGGACAGAGCCATCCAGAAATTGTTTCACAGAGCTAAATGT 1080
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1021 GCCTGTTGTGTCATAGGACGTGGACAGAGCCATCCAGAAATTGTTTCACAGAGCTAAATGT 1080

Qy 1081 GTTGTCTCAGACCTAATACAGAATCTACTGGTTTACATCACCATGTCTCTGGTTCTATAACA 1140
 |||
 Db 1081 GTTGTCTCAGACCTAATACAGAATCTACTGGTTTACATCACCATGTCTCTGGTTCTATAACA 1140

Qy 1141 TTCGAACAGACCCCTGGAGGATCAACACATATGACGGCTGATCTCAAAGGATTTAACGTT 1200
 |||
 Db 1141 TTCGAACAGACCCCTGGAGGATCAACACATATGACGGCTGATCTCAAAGGATTTAACGTT 1200

Qy 1201 AGTGAGGACTTGTCTCATCATCGTCATGGTGTGTCAGCTCCATGAATGGGGAGATATGTCC 1260
 |||
 Db 1201 AGTGAGGACTTGTCTCATCATCGTCATGGTGTGTCAGCTCCATGAATGGGGAGATATGTCC 1260

Qy 1261 CATGGCTGTCTACTCCTTAGGCAGAATGTACCATGGTCATGATGATGCTCATGACCCCAA 1320
 |||
 Db 1261 CATGGCTGTCTACTCCTTAGGCAGAATGTACCATGGTCATGATGATGCTCATGACCCCAA 1320

Qy 1321 AGACCTGGTGACCTTGGTGATGTTATAGATGATTCCCATGGCATCGTTCATTCAACTAGA 1380
 |||
 Db 1321 AGACCTGGTGACCTTGGTGATGTTATAGATGATTCCCATGGCATCGTTCATTCAACTAGA 1380

Qy 1381 ACCTTTGATCATCTTAATGTTGAAGATCTTAACGCACGTTCCCTTGTGATTATGCAGGGC 1440
 |||
 Db 1381 ACCTTTGATCATCTTAATGTTGAAGATCTTAACGCACGTTCCCTTGTGATTATGCAGGGC 1440

Qy 1441 GGACATGAGGTCTGAGAGTGAGAGGGTTGCTTGCTGTGTTATAGGACGGGCA 1491
 |||
 Db 1441 GGACATGAGGTCTGAGAGTGAGAGGGTTGCTTGCTGTGTTATAGGACGGGCA 1491

RESULT 2

BD268170

LOCUS BD268170 1611 bp DNA linear PAT 17-JUL-2003

DEFINITION Serine protease inhibitors.

ACCESSION BD268170

VERSION BD268170.1 GI:33077938

KEYWORDS JP 2002534063-A/2.

SOURCE unidentified

ORGANISM unidentified

unclassified.

REFERENCE 1 (bases 1 to 1611)

AUTHORS Scotti,P.D., Dearing,S.C., Greenwood,D.R. and Newcomb,R.D.

TITLE Serine protease inhibitors

JOURNAL Patent: JP 2002534063-A 2 15-OCT-2002;

THE HORTICULTURE AND FOOD RESEARCH INSTITUTE OF NEW ZEALAND LTD

COMMENT OS Shellfish

PN JP 2002534063-A/2

PD 15-OCT-2002

PF 23-DEC-1999 JP 2000591076

PR 23-DEC-1998 NZ 333568,23-JUL-1999 NZ 336906 PI

PAUL DOUGLAS SCOTTI,SALLY CAROLINE DEARING,DAVID ROGER PI

GREENWOOD,

PI RICHARD DAVID NEWCOMB

PC C12N15/09,A23L1/305,A61K38/00,A61P7/04,A61P43/00,C07K1/14, PC
 C07K14/435,

PC C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N9/99// (C12N9/99,C12R1:

PC 91),

```
FEATURES                                Location/Qualifiers
     source                               1. .1611
                                     /organism="unidentified"
                                     /mol_type="genomic DNA"
                                     /db_xref="taxon:32644"
```

Query Match 100.0%; Score 1490.6; DB 6; Length 1611;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1491; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

Db	601	 GATAAAACTGAGCATTATGCCCATTTGTGACGTAAGATCTAATACACACCAACCAAAGGCT	660
Qy	661	CTTCATCATCATGTCCACGGAACCATCGATTTCAAACAAGTTGGTTATGGTGACCTTGAA	720
Db	661	CTTCATCATCATGTCCACGGAACCATCGATTTCAAACAAGTTGGTTATGGTGACCTTGAA	720
Qy	721	GTGTCCTACCATTTAGAGGGATTTAATGTAAGTGATGACCACAAAGATCATCTCCATGAC	780
Db	721	GTGTCCTACCATTTAGAGGGATTTAATGTAAGTGATGACCACAAAGATCATCTCCATGAC	780
Qy	781	GTACAGATCTACGCCAACGGTGACCTGACCAGTGGATGTGATAACCTCGGTGCTAAATAT	840
Db	781	GTACAGATCTACGCCAACGGTGACCTGACCAGTGGATGTGATAACCTCGGTGCTAAATAT	840
Qy	841	GATCCTCATGAAGATTACCACAGTGAGTTGGGTGATCTAGGAGATATTCACGATGATGAC	900
Db	841	GATCCTCATGAAGATTACCACAGTGAGTTGGGTGATCTAGGAGATATTCACGATGATGAC	900
Qy	901	CATGGCGTTGTCAATGAAAGCCACAGATATTCTGGATCAATATCTTCGGTGATGACAGT	960
Db	901	CATGGCGTTGTCAATGAAAGCCACAGATATTCTGGATCAATATCTTCGGTGATGACAGT	960
Qy	961	GTCCTGGGACGTTCTATTGCCATTACCAAAGAGACCATCTTCATAAAAGTGCCAAAATT	1020
Db	961	GTCCTGGGACGTTCTATTGCCATTACCAAAGAGACCATCTTCATAAAAGTGCCAAAATT	1020
Qy	1021	GCCTGTTGTGTCATAGGACGTGGACAGAGCCATCCAGAAATTGTTACAGAGCTAAATGT	1080
Db	1021	GCCTGTTGTGTCATAGGACGTGGACAGAGCCATCCAGAAATTGTTACAGAGCTAAATGT	1080
Qy	1081	GTTGTCAGACCTAATACAGAATCTACTGGTTTACATCACCATGTCTCTGGTTCTATAACA	1140
Db	1081	GTTGTCAGACCTAATACAGAATCTACTGGTTTACATCACCATGTCTCTGGTTCTATAACA	1140
Qy	1141	TTCGAACAGACCCCTGGAGGATCAACACATATGACGGCTGATCTCAAAGGATTTAACGTT	1200
Db	1141	TTCGAACAGACCCCTGGAGGATCAACACATATGACGGCTGATCTCAAAGGATTTAACGTT	1200
Qy	1201	AGTGAGGACTTGTACATCATCGTCATGGTGTGCAGCTCCATGAATGGGGAGATATGTCC	1260
Db	1201	AGTGAGGACTTGTACATCATCGTCATGGTGTGCAGCTCCATGAATGGGGAGATATGTCC	1260
Qy	1261	CATGGCTGTCACTCCTTAGGCAGAATGTACCATGGTCATGATGATGCTCATGACCCCCAAA	1320
Db	1261	CATGGCTGTCACTCCTTAGGCAGAATGTACCATGGTCATGATGATGCTCATGACCCCCAAA	1320
Qy	1321	AGACCTGGTGACCTTGGTGATGTTATAGATGATTCCCATGGCATCGTTCATTCAACTAGA	1380
Db	1321	AGACCTGGTGACCTTGGTGATGTTATAGATGATTCCCATGGCATCGTTCATTCAACTAGA	1380
Qy	1381	ACCTTTGATCATCTTAATGTTGAAGATCTTAACGCACGTTCCCTTGTGATTATGCAGGGC	1440
Db	1381	ACCTTTGATCATCTTAATGTTGAAGATCTTAACGCACGTTCCCTTGTGATTATGCAGGGC	1440
Qy	1441	GGACATGAGGTCGAGAGTGAGAGGGTTGCTTGCTGTGTTATAGGACGGGCA	1491

Db 1441 GGACATGAGGTCGAGAGTGAGAGGGTTGCTTGCTGTGTTATAGGACGGGCA 1491

Db 814 GTGTCCTACCATTTAGAGGGATTTAATGTAAGTGATGACCACAAAGATCATCTCCATGAC 873
 Qy 781 GTACAGATCTACGCCAACGGTGACCTGACCAGTGGATGTGATAACCTCGGTGCTAAATAT 840
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 874 GTACAGATCTACGCCAACGGTGACCTGACCAGTGGATGTGATAACCTCGGTGCTAAATAT 933
 Qy 841 GATCCTCATGAAGATTACCACAGTGAGTTGGGTGATCTAGGAGATATTCACGATGATGAC 900
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 934 GATCCTCATGAAGATTACCACAGTGAGTTGGGTGATCTAGGAGATATTCACGATGATGAC 993
 Qy 901 CATGGCGTTGTCAATGAAAGCCACAGATATTCCTGGATCAATATCTTCGGTGATGACAGT 960
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 994 CATGGCGTTGTCAATGAAAGCCACAGATATTCCTGGATCAATATCTTCGGTGATGACAGT 1053
 Qy 961 GTCCTGGGACGTTCTATTGCCATTACCAAAGAGACCATCTTCATAAAAGTGCCAAAATT 1020
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1054 GTCCTGGGACGTTCTATTGCCATTACCAAAGAGACCATCTTCATAAAAGTGCCAAAATT 1113
 Qy 1021 GCCTGTTGTGTCATAGGACGTGGACAGAGCCATCCAGAAATTGTTACAGAGCTAAATGT 1080
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1114 GCCTGTTGTGTCATAGGACGTGGACAGAGCCATCCAGAAATTGTTACAGAGCTAAATGT 1173
 Qy 1081 GTTGTGACACCTAATACAGAATCTACTGGTTTACATCACCATGTCTCTGGTTCTATAACA 1140
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1174 GTTGTGACACCTAATACAGAATCTACTGGTTTACATCACCATGTCTCTGGTTCTATAACA 1233
 Qy 1141 TTCGAACAGACCCCTGGAGGATCAACACATATGACGGCTGATCTCAAAGGATTTAACGTT 1200
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1234 TTCGAACAGACCCCTGGAGGATCAACACATATGACGGCTGATCTCAAAGGATTTAACGTT 1293
 Qy 1201 AGTGAGGACTTGTCACATCATCGTCATGGTGTGCAGCTCCATGAATGGGGAGATATGTCC 1260
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1294 AGTGAGGACTTGTCACATCATCGTCATGGTGTGCAGCTCCATGAATGGGGAGATATGTCC 1353
 Qy 1261 CATGGCTGTCACTCCTTAGGCAGAATGTACCATGGTCATGATGATGCTCATGACCCCCAAA 1320
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1354 CATGGCTGTCACTCCTTAGGCAGAATGTACCATGGTCATGATGATGCTCATGACCCCCAAA 1413
 Qy 1321 AGACCTGGTGACCTTGGTGATGTTATAGATGATTCCCATGGCATCGTTCATTCAACTAGA 1380
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1414 AGACCTGGTGACCTTGGTGATGTTATAGATGATTCCCATGGCATCGTTCATTCAACTAGA 1473
 Qy 1381 ACCTTTGATCATCTTAATGTTGAAGATCTTAACGCACGTTCCCTTGTGATTATGCAGGGC 1440
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1474 ACCTTTGATCATCTTAATGTTGAAGATCTTAACGCACGTTCCCTTGTGATTATGCAGGGC 1533
 Qy 1441 GGACATGAGGTCGAGAGTGAGAGGGTTGCTTGCTGTGTTATAGGACGGGCA 1491
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1534 GGACATGAGGTCGAGAGTGAGAGGGTTGCTTGCTGTGTTATAGGACGGGCA 1584

Database : N_Geneseq_23Sep04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%	Query					
No.	Score	Match	Length	DB	ID		Description	
1	1490.6	100.0	1491	3	AAA47150		Aaa47150 DNA encod	
2	1490.6	100.0	1611	3	AAA47151		Aaa47151 DNA encod	
3	128.4	8.6	606	8	AAD48291		Aad48291 Crassostr	
4	56.8	3.8	1083	5	AAS76745		Aas76745 DNA encod	
5	54.4	3.6	2000	8	ADA71938		Ada71938 Rice gene	
6	52.8	3.5	110000	12	ADO34927_1		Continuation (2 of	
7	52.6	3.5	583	4	AAI23356		Aai23356 Probe #13	
8	52.6	3.5	583	4	ABA68463		Aba68463 Human foe	
9	52.6	3.5	583	4	AAI48680		Aai48680 Probe #17	
10	52.6	3.5	583	4	ABA50512		Aba50512 Human bre	

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	% Query						Description
	No.	Score	Match	Length	DB	ID	
	1	54.4	3.6	480	4	US-09-248-796A-6301	Sequence 6301, Ap
	2	51.2	3.4	291	4	US-09-248-796A-6300	Sequence 6300, Ap
	3	46.4	3.1	549	4	US-09-248-796A-3913	Sequence 3913, Ap
c	4	44.8	3.0	5340	4	US-09-627-122-21	Sequence 21, Appl
	5	42.4	2.8	2518	3	US-09-433-699-3	Sequence 3, Appli
	6	42.4	2.8	10304	4	US-09-627-465B-1	Sequence 1, Appli
	7	42	2.8	496	1	US-08-263-413-23	Sequence 23, Appl
	8	42	2.8	500	1	US-08-263-413-22	Sequence 22, Appl
	9	42	2.8	675	1	US-07-807-043B-2	Sequence 2, Appli
10	42	2.8	675	1	US-08-299-849B-2	Sequence 2, Appli	
11	42	2.8	675	2	US-08-142-368A-2	Sequence 2, Appli	
12	42	2.8	675	3	US-08-967-727-2	Sequence 2, Appli	
13	42	2.8	675	3	US-08-037-230D-2	Sequence 2, Appli	
14	42	2.8	675	4	US-09-583-850-2	Sequence 2, Appli	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result			%				ID	Description
	No.	Score	Match	Length	DB			
	1	52.6	3.5	583	9	US-09-864-761-20772		Sequence 20772, A
	2	52.6	3.5	1959	9	US-09-864-761-4012		Sequence 4012, Ap
c	3	49.2	3.3	327	9	US-09-864-761-28059		Sequence 28059, A
	4	49	3.3	676	17	US-10-437-963-44631		Sequence 44631, A
c	5	48.2	3.2	744802	15	US-10-292-798-1369		Sequence 1369, Ap
	6	47.8	3.2	1168	15	US-10-017-161-2179		Sequence 2179, Ap
	7	47.8	3.2	1168	15	US-10-292-798-1825		Sequence 1825, Ap
	8	47	3.2	1631	15	US-10-369-493-36458		Sequence 36458, A
	9	46.8	3.1	728	17	US-10-767-795-5840		Sequence 5840, Ap
	10	46.6	3.1	717	18	US-10-425-115-15020		Sequence 15020, A
c	11	46.2	3.1	456	9	US-09-864-761-11468		Sequence 11468, A
	12	45.8	3.1	493	17	US-10-767-701-31233		Sequence 31233, A
c	13	45.8	3.1	785	15	US-10-029-386-22627		Sequence 22627, A
c	14	45	3.0	506	15	US-10-029-386-20619		Sequence 20619, A
c	15	44.8	3.0	58985	10	US-09-901-152-3		Sequence 3, Appli
c	16	44.8	3.0	143601	10	US-09-855-824-3		Sequence 3, Appli
	17	44.4	3.0	1028	18	US-10-739-930-4488		Sequence 4488, Ap
	18	44.2	3.0	574	9	US-09-864-761-228		Sequence 228, App
	19	44.2	3.0	669	9	US-09-864-761-17051		Sequence 17051, A
	20	44.2	3.0	926	18	US-10-425-115-54567		Sequence 54567, A